DOCUMENT RESUME

ED 298 145 TM 012 232

AUTHOR Marsh, Herbert W.; Grayson, David
TITLE Public/Catholic Differences in the High School and Beyond Data: A Multi-Group Structural Equation Modelling Approach to Testing Mean Differences.
PUB DATE 88
NOTE 43p.; Some pages are marginally legible.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Academic Achievement; Academic Aspiration; *Achievement Tests; *Catholic Schools; Comparative Analysis; Differences; *Grade 12; High Schools; High School Students; Latent Trait Theory; Outcomes of Education; *Public Schools; Religious Education; School Effectiveness; *Scores
IDENTIFIERS *High School and Beyond (NCES); Latent Structure Analysis; LISREL Computer Program; *Structural Equation Modeling Approach

ABSTRACT

Previous research with the large, nationally representative High School and Beyond data has compared senior year achievement test scores for public and Catholic high school students after controlling for background variables and sophomore year test scores. These analyses, however, were based on traditional applications of multiple regression with its implausible assumptions that variables are measured without error and that residuals are uncorrelated. The present study demonstrates tests for mean differences on latent constructs using the LISREL approach to multi-group structural equation modelling for this substantively important issue. Data on 10,507 students from 853 public and 80 Catholic high schools were analyzed. Public/Catholic differences, even after controlling for background and sophomore outcomes, favored Catholic high school students on senior year outcomes (achievement, educational aspirations, and academic course selection) and subsequent college attendance. These public/Catholic differences were similar for students differing in race, socioeconomic status, and initial ability. However, public/Catholic differences in achievement, educational aspirations, and college attendance were apparently mediated by the academic orientation of course selection. The flexibility, advantages, and limitations of this multi-group structural equation modelling approach are discussed. Six tables and multiple figures and equations conclude the document. (Author/TJH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED298145

Public/Catholic Differences in the High School and Beyond Data: A Multigroup Structural Equation Modelling Approach to Testing Mean Differences

Herbert W. Marsh and David Grayson

University of Sydney

Running Head: Latent Mean Differences

The data used in this paper were made available by the Inter-university Consortium for Political and Social Research.

PRINTER NOTE: In order to avoid confusion, all Greek notation is presented in < >. For present purposes the following convention was used: <LX> = Lambda-X (Λ_x), <LY> = Lambda-Y (Λ_y), <BE> = Beta (β), <GA> = Gamma (Γ), <PH> = Phi (Φ), <PS> = Psi (Ψ), <TE> = Theta-Epsilon (Θ_ε), <TD> = Theta-delta (Θ_δ), <al> = alpha ($\underset{\sim}{\alpha}$), <xi> = xi ($\underset{\sim}{\xi}$), <ze> = zeta ($\underset{\sim}{\zeta}$), <et> = Eta ($\underset{\sim}{\eta}$), <mu> = mu ($\underset{\sim}{\mu}$).

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

HERBERT W. MARSH

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

TM012 232

BEST COPY AVAILABLE

Public/Catholic Differences in the High School and Beyond Data: A Multi-group Structural Equation Modelling Approach to Testing Mean Differences

ABSTRACT

Previous research with the large, nationally representative High School and Beyond (HSB) data has compared senior year achievement test scores for public and Catholic high school students after controlling for background variables and sophomore year test scores. These analyses, however, were based on traditional applications of multiple regression with its implausible assumptions that variables are measured without error and that residuals are uncorrelated. The present study demonstrates tests for mean differences on latent constructs using the LISREL approach to multi-group structural equation modelling for this substantively important issue. Public/Catholic differences, even after controlling for background and sophomore outcomes, favored Catholic high school students on senior year outcomes (achievement, educational aspirations, and academic course selection) and subsequent college attendance. These public/Catholic differences were similar for students differing in race, SES, and initial ability. Public/Catholic differences in achievement, educational aspirations and college attendance were, however, apparently mediated by the academic orientation of course selection. The flexibility and advantages -- but also the limitations -- of this multi-group SEM approach are discussed.

Public/Catholic Differences in the High School and Beyond Data: A Multi-group Structural Equation Modelling Approach to Testing Mean Differences

The present investigation has two major purposes, a substantive purpose and a methodological purpose. The substantive purpose is to compare outcomes for students who have attended public and Catholic high schools based on the High School and Beyond (HSB) data. The methodological purpose is to demonstrate tests for mean differences in latent constructs using the LISREL approach to multi-group structural equation modelling (SEM). Because this methodological demonstration addresses complicated, substantive issues derived from previous research, the present study also provides an opportunity to examine the flexibility, strengths and limitations of this approach.

The Substantive Issue -- Public/Catholic Differences in the HSB Data

The large, nationally representative High School and Beyond (HSB) data base has stimulated considerable interest in academic achievement effects attributed to attending public and Catholic high schools in the U.S. Because this research is reviewed elsewhere (e.g., Alexander & Pallas, 1985; Hoffer, Greeley & Coleman, 1985; Jencks, 1985; Marsh, 1988; Willms, 1985; Wolfle, 1987), it is only summarized briefly. Using cross-sectional analyses based on just the first wave of HSB data (the 1980 cohorts of sophomores and seniors), Coleman, Hoffer and Kilgore (1981, 1982) concluded that Catholic school students learned more than public school students during their last two years of high school. This initiated heated debate (e.g., Goldberger & Cain, 1982), a flurry of reanalyses (Alexander & Pallas, 1983; Cain & Goldberger, 1983; Morgan, 1983) and rejoinders (e.g., Coleman & Hoffer, 1983). Whereas basic issues were unresolved, Jencks (1985) concluded that: "All parties to this debate agreed, however, that the 1980 data were not ideal for estimating the effect of Catholic schooling, since there was no completely satisfactory way of knowing what seniors in 1980 were like when they were sophomores in 1978, or what sophomores in 1980 would be like when they became seniors in 1982" (p. 128).

The availability of the second wave of (1982) data provided a much stronger basis for subsequent analyses. Jencks noted that (1985, p. 128): "One major purpose of the High School and Beyond (HSB) study was to assess the impact of different kinds of schooling on how much students learned in the last two years of schooling." Jencks (1985) also provided a summary and critique of the three initial major analyses of public/Catholic differences based on sophomore and senior responses by the 1980 sophomore cohort (Alexander & Pallas, 1985; Hoffer, Greeley & Coleman, 1985; Willms, 1985).

All three studies found that senior Catholic high school students outperformed public high school students on standardized achievement tests, even after controlling for family background and sophomore achievement. There were still, however, important unresolved differences in the three studies.

1. The three studies differed on the variables that were controlled in inferring sophomore-to-senior growth. All the studies corrected 1982 achievement test scores for at least some background variables and the matching 1980 achievement test score. Willms (1985) controlled 1982 test scores for other tests scores in addition to the matching test score. This approach is generally better in that a more broadly based set of control variables is more likely to control for pre-existing differences than is a single variable within the set of control variables. This approach also allows researchers to consider other senior year outcomes and post-secondary outcomes (available in the third wave of data) that do not have a matching 1980 outcome variable. Both Jencks (1985) and Hoffer et al. (1985) acknowledged the appropriateness of this approach but indicated that it made little difference in the results for achievement test scores.

2. Willms (1985) estimated a single regression equation from the total group covariance matrix, whereas Alexander and Pallas (1985) and Hoffer et al. (1985) estimated separate regression equations for the public and Catholic high school samples. Willms inferred public/Catholic differences from a dummy (dichotomous) variable that represented school type. The other two studies resulted in slightly different estimates of the effect depending on whether regression equations based on the public or Catholic school sample were used (see Jencks, 1985).

3. Alexander and Pallas (1985) corrected both the 1980 and 1982 test scores for internal consistency estimates of reliability (Heyns and Hilton, 1982). This approach assumes that "error" (i.e., random error and uniqueness) estimated from one administration of a test is uncorrelated with "error" estimated for a second administration of the test. When the same test is administered on two occasions, however, correlated uniquenesses (sometimes referred to as correlated errors) are likely. If, for example, the test-retest correlation exceeds the internal consistency estimates of reliability, then the "corrected" test-retest correlation would be greater than 1.0. Since there probably are correlated uniquenesses in the test scores, the Alexander and Pallas approach may seriously overestimate the correlation between 1980 and 1982 test scores (Hoffer, et al. 1985). On the other hand, when there are moderate amounts of measurement error, Cohen and

Cohen (1983) and Pedhazur (1982) and others have noted that the traditional regression equations can lead to grossly inaccurate results (also see Alexander & Pallas, 1985; Jencks, 1985). Jencks took solace in the observation that both approaches resulted in similar patterns of school-type effects though this observation was based only on achievement test scores. It should be noted, however, that two wrongs do not necessarily make a right even when they lead to similar conclusions. Furthermore, Hoeffer et al., noted that these two approaches did lead to very different conclusions about the presence -- and even the direction -- of interaction effects (see point 4). In summary, this problem was not satisfactorily addressed in any of the studies or in Jencks's critique of them.

4) Willms (1985) tested for interactions between the public/Catholic grouping variable and background characteristics by including cross-product terms in his single regression equation. Because these interaction terms added little to variance explained and were nearly always nonsignificant, his results suggested that public/Catholic differences were similar for different types of students. In the two-equation approach, tests of interactions were made by comparing the regression coefficients in each equation. Hoffer, et al (1985) reported that SES, dummy variables for being Black and being Hispanic, and initial achievement had more impact on achievement in public schools than in Catholic schools, though few of these interaction effects were statistically significant[1]. They interpretted this to mean that Catholic high schools more closely approximated a common-school ideal since initial disadvantages had less negative effects in Catholic high schools. In contrast, Alexander and Pallas (1985) found that after correcting for unreliability (see point 3) initially disadvantaged students were slightly better off in public schools, though few of these differences were statistically significant. Hoffer et al countered that the Alexander and Pallas results were an artifact of their inappropriate correction for unreliability. In his review, Jencks (1985) did not comment on the appropriateness of the Alexander and Pallas tests of interaction effects. Jencks concluded that whereas there was little convincing evidence for the existence of interaction effects, these tests were not sufficiently powerful to conclude that they did not exist. Kenny and Judd (1984) and others have commented more generally on the problems of trying to infer interaction effects from fallible measures -- even when a latent construct approach is used.

5. Throughout the HSB studies of public/Catholic differences, high school track has played a controversial role. Researchers have consistently

found that Catholic high school students are more likely to be in academic tracks during their sophomore year. In early analyses conducted before the second wave of data was available, Coleman et al. (1982) argued that this represented a school-policy difference that produced achievement differences whereas others suggested the effect of track should be controlled as a proxy of initial ability differences. When the second wave of HSB became available Hoffer, et al. (1985) found that school-average measures of the proportion of students in academic tracks during their sophomore year accounted for some -- but far from all -- of the public/Catholic differences in senior achievement test scores. They concluded that their results provided at least some support for their earlier claims about track effects. In contrast, using their correction for unreliability Alexander and Pallas (1985) found that controlling for track had almost no effect on public/Catholic differences in senior achievement test scores beyond what could be explained by background variables and sophomore test scores. They concluded that there was no support for Coleman et al.'s earlier claims. Willms (1985) and Jencks (1985) did not emphasize the effect of school track. None of these studies considered the effects of track placement in the senior year, nor whether public/Catholic differences in this variable were stable over the last two years of high school.

Marsh (1988) further analyzed public/Catholic differences on the basis of three waves of HSB data instead of just the two waves previously available. He, as did Willms (1985), used a dummy variable in a single regression equation to assess public/Catholic differences after testing for many possible interaction effects. Two particularly important variables that he considered were not available previously: (a) the number of academically oriented courses completed by students as determined by an evaluation of their actual high school transcripts by HSB staff and (b) college attendance in the first two years after high school. The public/Catholic difference in academic track actually grew larger during the last two years of high school. The largest Catholic advantages -- after controlling for background and sophomore outcomes -- were for the new coursework selection variables. There was also a Catholic high school advantage in college attendance after controlling for background and sophomore outcomes, though this effect was largely mediated through senior outcome variables. Furthermore, controlling for track and the academic course selection in the senior year in addition to background and sophomore outcome variables, eliminated all statistically significant public/Catholic differences in the remaining senior and post-

secondary outcomes. Marsh interpretted these findings to suggest that the observed public/Catholic differences were largely mediated by differences in the academic orientation of course selection, findings that supported the Hoffer, et. al. (1985) proposal. Marsh also found that public/Catholic differences were very similar for single-sex and coed Catholic schools and that public/Catholic differences were very small -- mostly nonsignificant -- for affective variables such as locus of control, esteem, academic self-concept. Although Marsh provided new evidence about public/Catholic differences based on the HSB data, his study suffered limitations of earlier research that are inherent in the traditional use of multiple regression.

Individual test scores in specific subjects have been the basis of most of public/Catholic comparisons with the HSB data. Because there is only a single total score for each test, there is no fully adequate way to estimate reliability. Because they had access to item level data, Heyns and Hilton (1982) provided KR-20 reliability estimates for the tests administered in the first year of the study. Whereas this is useful information, correcting correlations for these estimates as did Alexander and Pallas (1985), is likely to provide an inflated estimate of the test-retest correlation for matching sophomore and senior tests. Heyns and Hilton (1982) also factor analyzed the set of sophomore tests. They concluded that there was "little empirical basis for hypothesizing more than two factors in the battery" (p. 93). The two factors were readily identifiable as a verbal and a math factor. Because there are multiple indicators of each of these latent constructs, it is possible to obtain estimates of the error/uniqueness associated with each test without item-level data and to test for the existence of correlated uniquenesses across the two testing occasions. Implicit in this approach is the not inconsequential assumption that the small amount of uniqueness (or specific variance) associated with some of the tests (see Heyns & Hilton, 1982) is not of interest.

All four studies examined here as well as others known to the authors were based on traditional applications of multiple regression using ordinary least squares estimates. Although widely used, this approach imposes apparently unrealistic assumptions such as (a) variables are measured without error and (b) residuals are uncorrelated. Particularly when there are multiple indicators of most of the underlying constructs, SEM provides a more powerful approach that does not impose these unrealistic assumptions. The purpose of the present investigation is to further analyze public/Catholic differences using the multi-group LISREL (Joreskog & Sorbom, 1981) approach to

testing differences in latent construct means as described by Sorbom (1982; also see Hanna and Lei, 1985; Joreskog and Sorbom, 1981; and Sorbom, 1974).

Methods

Sample and Variables

Data for the present investigation are based on responses by the sophomore cohort of the HSB study. A detailed description of this data base is available in the user's manual produced by the National Center for Educational Statistics (NCES, 1986). The data file includes variables collected in 1980 when respondents were sophomores, in 1982 when respondents were seniors, and in 1984 two years after the normal time of high school graduation. The sophomore cohort initially involved a two-stage probability sample of 1,015 high schools and approximately 36 sophomores within each of these schools. The second follow-up consisted of a probability sample of 14,825 of the original sample. For present purposes, students were selected from the second follow-up who: (a) attended a public or Catholic high school (private school students were excluded) and (b) attended the same school in 1980 and 1982 (students who had the same school identification number both years, had not dropped out, had not transfered to another school, and had not already graduated). This left a total of 10,507 students from 853 public schools and 80 Catholic schools.

Responses in the present analysis were weighted so as to hold constant the total sample size but to take into account the disproportionate sampling of specified subgroups -- particularly the over-sampling of Catholic high school students -- in the HSB design (NCES, 1986, Table 3.5-1). The original unweighted and subsequently weighted sample sizes were: 8175 and 9744 (public school); 2332 and 763 (Catholic school). Because of the cluster sampling in the HSB study, standard errors based on the assumption of simple random sampling substantially underestimate the sampling variability in summary statistics and distort tests of statistical significance. In order to compensate for this bias, the weight for each respondent was divided by the estimated design effect of 2.40 (NCES, 1986, Table 3.6-5), reducing the nominal sample size from 10,507 to 10,507/2.4=4,378. (This reduction in nominal sample size has no effect at all on cell means and parameter estimates; it only affects the df used in tests of statistical significance.) Moment matrices were constructed for each group separately using pair-wise deletion for missing data. The weighted number of cases for each variable varied from 3656 to the maximum of 4,378, and the minimum pairwise number of cases was 3232. For purposes of testing statistical significance, sample sizes of 300 and 3,700 were used for the Catholic and

public schools respectively.

The 27 variables selected for consideration (see Appendix 1 and Table 1) were classified as background variables, or as sophomore, senior, and post-secondary outcome variables. Background variables were two indicators of social economic status (SES) and two variables that reflected race/ethnicity. A total of 23 outcome variables were selected to represent potentially important influences of school-type; 20 (10 pairs) were matching measures collected in both sophomore and senior years, 1 was a senior outcome that had no matching sophomore outcome, and 2 were multiple indicators of post-secondary education. These outcome measures included standardized achievement tests, educational aspirations, academic track, number of academic credits measured in the sophomore and/or senior years, and college attendance during the two years after the normal graduation from high school.

Insert Table 1 About Here

The Application of the LISREL Approach to SEM of Group Differences.

Weaknesses of the traditional use of multiple regression for estimating path coefficients, particularly for longitudinal data, are well known (e.g., Joreskog & Sorbom, 1981; Kenny, 1979; Long, 1983a, 1983b; McDonald, 1985; Pedhazur, 1982; Rogossa, 1979) and so are not reviewed here in detail. Perhaps the most serious weakness in the traditional multiple regression approach is the assumption that the single score used to infer each construct is measured without error. Particularly when multiple indicators of the inferred constructs are available, SEM using statistical packages such as LISREL provide important advantages.

The A Priori Measurement and Structural Models.

Although parameters for the entire model are estimated simultaneously, the model can be logically separated into measurement and structural models. The measurement model contains estimates of the relations between each latent construct and its multiple indicators (i.e., factor loadings) and error/uniquenesses associated with each measured variable. The structural model contains estimates of causal relations between the latent constructs (i.e., path coefficients) that are corrected for measurement error.

The measurement model. The a priori measurement model is summarized in Table 1. The latent constructs (<et>$_1$ to <et>$_{13}$) associated with each measured variable (y1 to y27) are presented in Table 1. Most of the latent constructs are associated with two or more measured variables and most of the measured variables are associated with only one latent construct. The factor loading of the first indicator of each construct was fixed to be 1 in

order to set the scale of the latent construct.

Many of the measured variables associated with the sophomore and senior constructs were actually the same variable administered on two different occasions. It was posited a priori that each of the uniquenesses for measured variables from the sophomore year (the multiple indicators of math and verbal achievement, and of educational aspirations -- y5-y13 in Table 1) was correlated with the uniqueness for the matching senior indicator (y15-y23). In preliminary analyses, 8 of these 9 correlated uniquenesses -- all but the one relating y5 and y15 -- were statistically significant and these 8 correlated uniquenesses were retained in subsequent analyses. The scales for each of the latent constructs were set by fixing the factor loading for the first measured variable to be 1.0. For all single-indicator latent constructs, the uniqueness of the single indicator was fixed to be zero.

The latent constructs are grouped into clusters representing background constructs, sophomore constructs, senior constructs, and the single post-secondary construct. For the first three clusters that contain multiple constructs, the latent constructs within the same cluster were posited to be correlated.

Structural model. In the initial, a priori model, latent constructs at each level were posited to effect all subsequent constructs: (a) background constructs affected sophomore, senior, and post-secondary outcomes, (b) sophomore outcomes affected senior and post-secondary constructs, and (c) senior constructs affected the post-secondary construct. In subsequent models variations of this initial structural model were considered in order to test substantive issues.

Insert Figure 1 About Here

Factorial invariance. Tests of factorial invariance require that any one, any set, or all parameter estimates in one group are the same across multiple groups -- in this case the public and Catholic high school samples. The advantages of CFA approach to factorial invariance over the comparison of solutions based on exploratory factor analyses are well-known (Alwin & Jackson, 1981; Joreskog & Sorbom, 1981; Marsh, 1987; Marsh & Hocevar, 1985) and so well not be considered here in detail. Whereas tests of the invariance of any or all parameters can be tested, the minimum condition of factorial invariance is frequently taken to be the invariance of the factor loadings (e.g., Marsh & Hocevar, 1985). Thus, the fit of one model that requires all factor loadings to be the same is compared with the fit of another model that does not require this invariance (see earlier discussion

of goodness of fit). Byrne and Shavelson (in press) and Byrne, Shavelson and Muthen (1987) have noted, however, that further tests are meaningful even when strict tests of factorial invariance are not supported. In these applications, however, differences in the factor loadings for the two groups were very small.

The analysis of structured means. Following Sorbom (1982; Joreskog and Sorbom, 1981) the analysis of mean structures was carried out on an augmented (i.e., a constant variable 1 was added) moment matrix. The input to LISREL comprised: (a) two 28x28 correlation matrices including one row of zeros corresponding to the constant variable in addition to the 27 variables listed in Table 1; (b) two vectors of 28 means in which the constant term was assigned a mean of 1; and (c) two vectors of standard deviations in which the constant term was assigned a value of zero. LISREL constructed separate moment matrices for the public and Catholic school samples.

The mathematical translation of Sorbom's approach to the present investigation and the input to the LISREL V program used to estimate the parameters is presented in the Appendix 2. Briefly: (a) the original 27 measured variables are specified to be ys and the associated 13 latent constructs were specified to be <et>s; (b) the "constant" variable was given a fixed-X specification; (c) the factor loading <LX> matrix was augmented to accomodate the constant variable by adding one column (the 14th) that was used to represent measurement intercepts <mu>s; (d) the <PS> matrix was augmented to accomodate the constant variables by adding an extra row (the 14th) of zeros; (e) mean differences in latent constructs were estimated in <GA> by fixing all the parameter estimates <al>s to be zero in one group and freely estimating the these parameters in the second group. Since the measurement intercepts and latent means cannot be identified simultaneously, absolute mean estimates are not possible. Latent mean differences between groups are estimated by fixing the latent means in one group to be zero, freely estimating the latent means in the second group, and testing whether the latent mean estimates in the second group differ significantly from zero. Comparing the X^2 for two nested models that differ only due constraining latent mean differences to be zero (i.e., fixing parameter estimates in <GA> to be zero in both groups) provides a multivariate omnibus test of latent mean differences.

Diagrams used to represent models of latent mean differences. As described in Appendix 2, we have endeavored to represent models of latent mean differences so as to highlight important features. To illustrate this

convention, consider model M2 (figure 1). Ovals represent clusters of latent constructs that do not causally influence each other: (1) the 3 background constructs; (2) the 4 sophomore constructs; (3) the 5 senior constructs; and (4) the 1 post-secondary construct. For clarity we have totally omitted the measurement components of these constructs. The only parameters on the graph are those free to vary between the two groups (i.e., public and Catholic). Single-headed arrows show putative causal links, whereas double-headed arrows represent covariances. All constructs are assumed to causally influence those to the right, although only lag-1 arrows are shown. The triangle represents the "constant" variable used by LISREL to estimate latent mean differences. Thus, in model M2 (figure 1), $\langle a1 \rangle_{13}$ represents the mean difference in post-secondary attendance (see Table 1) after adjusting for background, sophomore, and all senior consstructs. In model M8, only the background and sophomore constructs are adjusted for, whereas in model M9 $\langle a1 \rangle_{13}$ represents the sector difference on post-secondary attendance adjusting for background, sophomore and just two of the senior outcomes. (The substantive basis of these models is decribed in detail as part of the presentation of the results.)

Goodness of fit.

An important, unresolved issue in SEM is the assessment of goodness of fit. On the basis of theory and previous research, the researcher typically posits a set of alternative models designed to explain relations among the measured variables. To the extent that the hypothesized model is able to fit the observed data, there is support for the model. The problem of goodness of fit is how to decide whether the predicted and observed results are sufficiently alike to warrant support for a model. Whereas X^2 values can be used to test whether these differences are statistically significant, there is a growing recognition of the inappropriateness of the classical hypothesis testing approach. Because restricted models are only designed to approximate reality, all such models are a priori false and will be shown to be false if tested with a sufficiently large sample size (Cudeck & Browne, 1983; Marsh, Balla & McDonald, 1988; Marsh, McDonald & Balla, 1988; McDonald, 1985). This problem is particularly obvious in studies like the present one in which sample sizes are so large that even trivial differences will result in large, statistically significant X^2s. Model selection must be based on a subjective combination of substantive issues, inspection of parameter values, goodness of fit, model parsimony, and a comparison of the performances of competing models. A variety of fit indices have been derived

to aid in this process such as the X^2/df ratio, the Tucker-Lewis Index (TLI; Tucker & Lewis, 1957), the Bentler-Bonett Index (BBI; Bentler & Bonett, 1980) that are considered here. In simulation studies of these and other indices Marsh, Balla and McDonald (1988) and Marsh, McDonald, and Balla (1988) found that the X^2/df and the TLI imposed apparently appropriate penalty functions for the inclusion of additional variables that controlled for capitalizing on chance, whereas the TLI was also relatively independent of sample size.

Results and Discussion

Tests of Alternative Models.

The initial model considered here (M1 in Table 2) was based on the assumption that all parameters except the latent construct means were invariant across the two groups. Due primarily to the large sample size, the X^2 for this model (1450 with df=632) is statistically significant, but the fit indices (e.g., TLI=.987) suggest that the fit is reasonable. Inspection of the standard deviations for the measured variables (Table 1) indicates, however, that group variances for the first four variables -- those associated with SES and race -- differ substantially for public and Catholic high school students. Not surprisingly, the modification indices provided by LISREL suggested that allowing the variances and covariances associated with these three latent constructs (SES, Black, Hisp) to differ across the two groups would improve the fit. In model M2 (Figure 1) these 6 parameters were freed, that is were not held invariant across the two groups. [Note: In subsequent discussion we use the expression "freeing parameters" to mean estimating parameters separately for each group.] The change in X^2 (107) was substantial in relation to the change in df (6), indicating the superiority of model M2. It should be noted, however, that the changes in fit indicators (e.g., TLIs of .989 vs. .987) are very small.

Insert Table 2 About Here

In model M2, 160 parameter estimates were fixed to be invariant across the public and Catholic school data -- all but the 13 latent mean differences (i.e., latent construct intercepts) and the six parameters freed in M2. An omnibus test of the invariance of 133 of these parameters -- all but the intercept terms for the 27 measured variables -- was provided by model M3 in which all 133 parameters were freed. Despite the large sample size, the change in X^2 (149) was small in relation to the change in df (133) and not statistically significant. The TLI which takes into account the number of estimated parameters is marginally higher for M2 (.989) than for

M3 (.986). The overall goodness of fit of model M2 and this omnibus test of factorial invariance provide good support for Model M2.

Model M4 provided a multivariate test of adjusted group differences on the 6 (adjusted) latent group means reflecting the senior year and post-secondary outcome variables. Model M4 differed from model M2 in that mean differences on these 6 latent constructs were fixed to be 0. The change in X^2 (147) in relation to the change in df (6) was large and statistically significant.

Parameter Estimates For the Selected Model M2.

On the basis of tests summarized in Table 2 and an inspection of the parameter estimates, model M2 was selected as the most appropriate model. Estimated factor loadings (Table 3) suggest that each of the latent constructs with multiple indicators is well defined. The correlated uniquenesses for matching measured variables assessed in the sophomore and senior years are substantial (Table 3). It is also interesting to note that factor loadings and uniquenesses associated with matching sophomore and senior variables are similar (Table 3). [Strict tests of the invariance of these parameters (model M5 in Table 3) failed, however, due primarily to factor loadings associated with the science and writing tests and the second indicator of educational aspirations.] Covariance terms relating the three background constructs -- SES, Black, and Hisp -- are all negative (Table 4). Residual covariances among the 4 sophomore outcome constructs and among the 5 senior outcome constructs are all positive (Table 4). The background constructs are substantially related to subsequent outcome measures (Table 4), though much of their effect on senior and post-secondary outcomes is mediated by sophomore outcomes.

Insert Tables 3 and 4 About Here

The most important parameters for present purposes are the public/Catholic differences in the latent construct means. Consistent with the design of the structural model, latent mean differences in: (a) the background constructs are unadjusted (i.e., adjusted and unadjusted means in Table 5 are the same), (b) the sophomore outcome constructs are adjusted for background constructs, (c) the senior outcome constructs are adjusted for background and sophomore constructs, and (d) the post-secondary construct are adjusted for background, sophomore, and senior constructs. Consistent with previous research with the HSB data, public/Catholic differences in background and sophomore outcomes are not interpretted as school-type effects, whereas public/Catholic differences in senior and post-secondary

outcomes that have been corrected for background and sophomore constructs are interpretted as school-type effects. All adjusted and unadjusted group differences on the latent means are presented in Table 5, but the emphasis in subsequent discussion is on the adjusted differences for senior and post-secondary outcomes.

Insert Table 5 About Here

For model M2, the selected model, public/Catholic differences in adjusted latent means are statistically significant for all 5 senior outcome constructs, but not for the post-secondary outcome. In model M2, the variances and covariances associated with the three background constructs were not held to be invariant across groups. Inspection of the latent mean differences for model M1 that forced these coefficients to be invariant (Table 5), however, shows that these restrictions had no effect on any of the latent mean differences (i.e., they are the same for the two models up to the third place beyond the decimal point that is normally presented by LISREL).

Model M3 did not constrain any of the parameters except the intercepts of the measured variables to be invariant (in order to make the model identified) and mean differences on such a model would generally be difficult to interpret. The results in Table 5, however, indicate that the latent mean differences in models M1 and M2 are very similar to the corresponding parameter estimates in model M3. This similarity is because all the remaining parameter estimates in the unconstrained model M3 are similar to those in models M1 and M2 that constrained all or most parameters to be invariant. These findings demonstrate that public/Catholic differences reported here are apparently robust with respect to restrictions placed on the invariance of parameter estimates in the two groups.

Interpretations of the latent mean differences suggest that Catholic high school students have better mathematics and verbal achievement, have higher educational aspirations, are more likely to be in the academic track, and take more academically related courses in their senior year of high school than do public school students even after controlling for SES, race, and sophomore outcome variables. Whereas public and Catholic high school students differ substantially on their (unadjusted) attendance of college during the first two years after high school, these differences are not significant after controlling for SES, race, sophomore and senior outcome variables.

Alternative Models Designed to Test Additional Substantive Issues.

An important advantage of the SEM approach to testing latent mean

difference is its flexibility. The selected model can be modified in a variety of ways to examine substantive issues that will often be idiosyncratic to a particular application. This flexibility is demonstrated with alternative models designed to address substantively important issues in the present investigation.

Public/Catholic differences on college attendance. The nonsignificant latent mean difference for the post-secondary construct (Table 5) should be interpreted cautiously, because this mean difference has been corrected for senior outcome constructs as well as for sophomore and background constructs. In the terminology of path analysis, it is the direct effect of the grouping variable on the post-secondary outcome construct. For the HSB data, it is also meaningful to consider the effect of the public/Catholic grouping variable that is mediated through the senior outcome constructs after controlling for background and sophomore constructs. Using the language of path analyses this effect would be the sum of the direct and indirect effects of the public/CAtholic grouping variable. For the HSB data, public/Catholic differences in background and sophomore outcomes are typically not considered to be school-type effects, and so it may not be meaningful to consider the effect of the public/Catholic grouping variable that are mediated through these constructs.

There are several ways in which model M2 could be modified to obtain the effect of school-type on university attendance correcting for background variables and sophomore outcomes. For present purposes, we posited an alternative structural model in which the post-secondary latent construct was merely correlated with five senior latent constructs (see M8 in Table 2 and Figure 1). This involved fixing the five path coefficients (in <BE>) leading from the senior constructs to the post-secondary outcome to be zero and estimating the corresponding 5 residual covariances (in <PS> subject to the constraint that these 5 parameter estimates were still invariant across groups. It is important to emphasize that M8 must necessarily provide the same fit as model M2 (so long as both models converge on a well-defined solution), but that adjusted mean difference in the post-secondary construct in model 8 is only adjusted for background and sophomore constructs. For model M8, the mean difference in the post-secondary construct is statistically significant (mean difference= -.169, SE=.047) and moderately large in comparison with the variance of the post-secondary construct (.757, Table 4). Based on the logic of the present application, we interpret this to mean that Catholic high school students attend college more than public

high school students even after controlling for background and sophomore outcomes, but that this difference is largely mediated by public/Catholic differences in senior year outcomes.

Interactions with initial student characteristics. As summarized earlier, an unresolved issue is whether public/Catholic differences vary according to SES, race, or initial ability level. The lack of resolution was due in part to the inability to appropriately control for measurement error. The previous comparison of models M2 and M3 provided a omnibus test of the invariance of 133 parameter estimates in public and Catholic high school samples. Because there was support for this invariance, the findings imply that the public/Catholic grouping variable does not interact significantly with these individual student characteristics.

The comparison of models M2 and M3 was based on so many parameter estimates, however, that it is possible that a few statistically significant and substantively important differences were overlooked. For this reason, more specific tests of these interactions and an inspection of the parameter estimates are considered in two additional models. Following previous research, further models were tested to determine the differential effects of SES, race, and initial ability in the public and Catholic samples. In model M6 the 30 path coefficients (in <BE>) leading from the three background constructs (SES, Black, and Hispanic) to the remaining 10 latent constructs were estimated independently for the two groups. In model M7, these path coefficients and the additional 24 path coefficients leading from the four sophomore constructs to the remaining 6 constructs were also estimated independently for the two groups. In each case, the changes in X^2 between model M2 and these additional models was small (23 and 36 respectively) in relation to the changes in df (30 and 54) and not statistically significant (see Table 2). Furthermore, the TLI index that adjusts for the number of estimated parameters was marginally poorer for these alternative models than for model M2.

Insert Table 6 About Here

The lack of significant differences between Models M2, M6 and M7 should, perhaps, preclude any further consideration of these interaction effects. Nevertheless, for purposes of illustration, parameters estimates for the public and Catholic samples that are most relevant to this issue are shown in Table 6. Only one public/Catholic difference was statistically significant at the nominal $p < .05$ level (none were significant at $p < .01$, and none of those not presented was statistically significant at either

[illegible] appear to be substantively small as well and [illegible]. The [illegible] differences, however, including [illegible] sophomore achievement [illegible], being black (negatively) [illegible] in the public sample than in the Catholic [illegible] the direction of these differences might be used to [illegible] contention that initially disadvantaged students are better off [illegible] Catholic high schools, such a conclusion may be unjustified. Specifically, because of the design of the HSB study, public/Catholic differences in sophomore outcomes are not interpretted as school type effects. It is possible, perhaps even likely, that low-SES and minority students who elect (or were selected) to attend Catholic high schools differed from those who did not. Because differences in the paths leading from the background and sophomore achievement constructs to senior and post-secondary outcome constructs are consistently nonsignificant and apparently trivial in size, there appears to be no support for the contention that public/Catholic differences vary depending on SES, race, or initial ability levels for the HSB data.

Process variables that mediate public/Catholic differences. Marsh (1988), suggested that public/Catholic differences in senior year achievement and subsequent college attendance may be mediated largely by the more academic orientation of courses selected by Catholic high school students. Two variables that he considered were included here: academic track and number of credits in academic courses. In support of this suggestion he demonstrated that public/Catholic differences were larger on these variables than other senior and post-secondary outcomes and that controlling for these senior outcomes rendered as nonsignificant the public/Catholic differences in the remaining senior and post-secondary outcomes.

In order to test this proposal, model M8 was altered so that latent mean differences in senior achievements, senior educational aspirations, and college attendance were adjusted for senior academic track and number of academic credits in addition to background and sophomore constructs (see Model M9 in Table 2 and Figure 1). This was accomplished by positing track and credits to directly effect the remaining senior and post-secondary outcomes instead of merely being correlated with them. Again, it should be noted that M9 must necessarily have the same fit as models M2 and M8, but differs in how mean differences in the latent constructs are adjusted. After adjusting for track and credits, mean differences were substantially smaller

and failed to reach statistically significance for senior math achievement (-.040), senior verbal achievement (-.056), senior educational aspirations (.005) and college attendance (-.046). Furthermore, when latent mean differences in these remaining four outcome constructs were fixed to be zero (model M10 in Table 2), the change in X^2 (5) was small in relation to the change in df (4) and not statistically significant. These results support the proposal that public/Catholic differences emphasized here and in previous research may be mediated largely by sector differences in the academic orientation of coursework. It is important to emphasize that the logic of this model still indicates that differences in senior achievement, educational aspirations and college attendance are legitimate school-type effects. In this respect, the analyses suggest what appears to be an important policy difference between public and Catholic high schools that may account for public/Catholic differences in other outcomes.

Summary and Implications

Analysis of covariance is routinely used in the comparison of nonequivalent groups despite its many implicit assumptions that are typically ignored and often untestable. The more general multiple regression approach to analysis of covariance provides important advantages, but still suffers important limitations that were discussed in relation to previous analyses of public and Catholic high school students in the HSB study. Methodologically, the present application of the SEM approach to multi-group comparisons of latent mean differences offers important advantages over traditional approaches to multiple regression used previously. In particular, the present approach:

1) tests the implicit factor structures used to form variables that have multiple indicators and the invariance of these factor structures across groups;

2) provides tests of and more appropriate corrections for measurement error and correlated uniquenesses;

3) tests for mean differences in latent constructs that have been appropriately corrected for measurement error; and

4) provides omnibus and specific tests of whether effects on the latent constructs attributable to grouping variables interact with other constructs.

The present application involved a moderately simple structure used to compare latent means in just two groups. The approach, however, can be easily extended to include a larger number of groups, factorial designs of multiple grouping variables, additional covariates or outcome measures,

fallible covariates and outcome measures that are parallel, tau-equivalent or congeneric, and provision for testing a wide variety of alternative structures that are idiosyncratic to a particular application (Sorbom, 1982; Joreskog & Sorbom, 1981). It is also important to note that many of the advantages of the SEM approach also apply to comparisons of randomly assigned groups in the same way that ANCOVA is used in both experimental and nonexperimental studies.

Despite the important advantages of the SEM approach to multi-group comparisons, there are also important limitations. First, there are practical limitations to the number of measured and latent variables and the number of groups that can be considered. These limitations are largely a function of the computer and software used in the analyses and further developments in the efficiency of the statistical procedures may resolve this problem. Second, the interpretation of latent mean differences becomes more problematic as the measurement and structural parameters become less invariant across groups. Most published examples, like the present application, have provided good support for the invariance of factor loadings and path coefficients relating covariates and outcome measures. The invariance of these path coefficients corresponds to the homogeneity of regression assumption in ANCOVA. When this assumption is not met, it is still possible to interpret group differences, but the size and even the direction of group differences will depend on values of one or more covariates (see Cohen & Cohen, 1985; Judd & Kenny, 1981, for a discussion of this issue in the multiple regression approach to ANCOVA that generalizes to the SEM approach). As of yet, there are apparently no widely agreed upon minimum conditions about what parameters must be invariant in order to compare latent mean differences or how robust existing approaches are to minor violations of these conditions.

A third limitation of the SEM approach to multiple-group comparisons of latent mean structures is that groups must be inferred from discrete, error-free grouping variables. Whereas this assumption is pluasible for a wide variety of grouping variables (e.g., sex, race, randomly assigned experimental and control groups) many grouping variables are continuous (e.g., age) and/or cannot be measured without error (SES). The application of the techniques described here requires continuous variables to be divided into discrete categories and measurement error to be ignored in assigning subjects to these discrete groups. There may be special conditions in which these pragmatic alternatives are acceptable, but in general they are not. A

viable alternative may be to conduct analyses on a total group covariance matrix in which grouping variables are represented as latent constructs. This approach provides for continuous, fallible grouping variables that are corrected for measurement error, but it imposes other important limitations. First, many of the tests of invariance that are so important in the multiple group comparisons are not easily performed and may not be testable. Second, in a related problem, nonlinear effects of latent constructs and interactions between latent constructs are not easily estimated with existing software. Kenny and Judd (1984) have described a procedure for estimating nonlinear and interactive effects of latent variables, but this procedure could not be implemented with LISREL (the authors used COSAN as described by McDonald, 1978), required use of generalized loss function instead of the maximum likelihood loss function, and did not provide any standard errors with which to evaluate parameter estimates. They concluded by noting that their procedure was "merely a beginning to developing a general approach to such estimation" (Kenny & Judd, 1984, p. 209). Interestingly, however, it is easy to test for interactions between discrete, error-free grouping variables (e.g., public and Catholic high school students) and a continuous latent construct based on fallible measures (e.g., SES) using the multiple-group approach as demonstrated in the present application. It would also be easy to test for interactions between two discrete, error-free grouping variables.

In summary, the SEM approach to multiple group comparisons of latent means has important advantages over traditional approaches to ANCOVA. The approach can readily be generalized to a wide variety of applications, but there are also important limitations. A particularly important limitation is that groups must be inferred on the basis of discrete, error-free grouping variables.

FOOTNOTES

1 -- Of the 24 coefficients presented by Hoffer, at al. (1985, Table 2.4), only five exceeded twice the reported standard error of the difference and only one exceeded three times the reported standard error. Assuming a design effect of 1.5, only one of the 24 differences is statistically significant at $p < .05$ and none is significant at $p < .01$. Also, the actual analysis presented by Hoeffer et al. included school-level means as well as individual student values for race, SES and sophomore achievement, and so these analyses may not be comparable to those presented by Alexander and Pallas (1985) that were based on just individual student values.

REFERENCES

Alexander, K. L., & Pallas, A. M. (1983). Private schools and public policy: New evidence on cognitive achievement in public and private school. Sociology of Education, 56, 170-82.

Alexander, K. L., & Pallas, A. M. (1985). School sector and cognitive performance. Sociology of Education, 58, 115-128.

Bentler, P. M. & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. Psychological Bulletin, 88, 588-606.

Cain, G. G., & Goldberger, A. S. (1983). Public and private schools revisited. Sociology of Education, 56, 208-218.

Cohen, J., & Cohen, P. (1983). Applied multiple regression/correlation analysis for the behavioral sciences (Second Edition). Hillsdale, NJ: Erlbaum.

Coleman, J. S., & Hoffer, T. (1983). Response to Taeuber-James, Cain-Goldberger and Morgan. Sociology of Education, 56, 219-234.

Coleman, J. S., Hoffer, T., & Kilgore, S. (1981). Public and private schools. Report to the National Center for Educational Statistics. Chicago: National Opinion Research Center.

Coleman, J. S., Hoffer, T., & Kilgore, S. (1982). Cognitive outcomes in public and private schools. Sociology of Education, 55, 66-76.

Coleman, J. S., Hoffer, T., & Kilgore, S. (1982). High school achievement. New York: Basic Books.

Cudeck, R., & Browne, M. W. (1983). Cross-validation of covariance structures. Multivariate Behavioral Research, 18, 147-167.

Goldberger, A. S., & Cain, G. G. (1982). The causal analysis of cognitive outcomes in the Coleman, Hoffer and Kilgore report. Sociology of Education, 55, 103-122.

Hanna, G. & Lei, H. (1985). A longitudinal analysis using the LISREL-model with structured means. Journal of Educational Statistics, 10, 161-169.

Heyns, B., & Hilton, T. L. (1982). The cognitive tests for High School and Beyond: An assessment. Sociology of Education, 55, 89-102.

Hoffer, T., Greeley, A. M., & Coleman, J. S. (1985). Achievement growth in public and Catholic schools. Sociology of Education, 58, 74-97.

James, L. R., Mulaik, S. A., & Brett, J. M. (1982). Causal analysis: Assumptions, models and data. Beverly Hills: Sage.

Jencks, C. (1985). How much do high school students learn. Sociology of Education, 58, 128-135.

Joreskog, K. G. & Sorbom, D. (1981). LISREL V: Analysis of Linear

Structural Relations By the Method of Maximum Likelihood. Chicago: International Educational Services.

Judd, C. M., & Kenny, D. A. (1981). Estimating the effects of social interventions. London: Cambridge University Press.

Kenny, D. A. (1979). Correlation and causality. New York: Wiley

Kenny, D. A., & Judd, C. M. (1984). Estimating the nonlinear and interactive effects of latent variables. Psychological Bulletin, 96, 201-210.

Long, K. S. (1983a) Confirmatory factor analysis: A preface to LISREL. Beverly Hills, CA: Sage.

Long, K. S. (1983b) Covariance structure models: An introduction to LISREL. Beverly Hills, CA: Sage.

Marsh, H. W., Balla, J. R. & McDonald, R. P. (1988). Goodness-of-fit indices in confirmatory factor analysis: The effect of sample size. Psychological Bulletin, xxx-xxx.

Marsh, H. W., McDonald, R. P., & Balla, J. R. (1988). Goodness of fit in confirmatory factor analysis: The effect of alternative models and sample size. (in review).

McDonald, R. P. (1978). A simple comprehensive model for the analysis of covariance structures. British Journal of Mathematical and Statistical Psychology, 31, 59-72.

McDonald, R. P. (1985). Factor analysis and related methods. Hillsdale, NJ: Lawrence Erlbaum.

Morgan, W. R. (1983). Learning and student life quality of public and private schools youth. Sociology of Education, 56, 187-202.

National Center for Educational Statistics (1986). High School and Beyond, 1980: Sophomore cohort second follow-up (1984), Data file user's manual. Ann Arbor, MI: Inter-university Consortium for Political and Social Research.

Pedhazur, E. J. (1982). Multiple regression in behavioral research (2nd Ed.). New York: Holt, Rinehart and Winston.

Rogossa, D. (1979). Causal models in longitudinal research: Rationale, formulation, and interpretation. In J. R. Nesselroade & P. B. Baltes (Ed.), Longitudinal research in the study of behavior and development (pp. 263-302). New York: Academic Press/

Sorbom, D. (1974). An general method for studying differences in factor means and factor structures between groups. British Journal of Mathematical and Statistical Psychology, 27, 229-239.

Sorbom, D. (1982). Structural equations models with structured means. In K. G. Joreskog & H. Wold (Eds.), Systems under indirect observation: Causality-Structure-Prediction. The Netherlands: North Holland Publishing.

Tucker, L. R. & Lewis, C. (1973). The reliability coefficient for maximum likelihood factor analysis. Psychometrica, 38, 1-10.

Willms, J. D. (1985). Catholic-school effects on academic achievement: New evidence from the High School and Beyond follow-up study. Sociology of Education, 58, 98-114.

Wolfle, L. M. (1987). Enduring cognitive effects of public and private schools. Educational Researcher, 16, 5-11.

Table 1

Public and Catholic Means and SDs For Measured Variables

Measured Varailes			Public School Mean	SD	Catholic School Mean	SD	Latent Construct
Background Variables							
1.	(y1)	SES-1	-0.043	0.994	0.509	0.932	SES <et>1
2.	(y2)	SES-2	-0.044	0.994	0.547	0.904	SES <et>1
3.	(y3)	Race--Black	0.018	1.021	-0.226	0.643	Black <et>2
4.	(y4)	Race--Hispanic	0.008	1.009	-0.099	0.874	Hisp <et>3
Sophomore Outcomes (based on 1980 data)							
5.	(y5)	Math-1	-0.031	1.004	0.362	0.871	Math <et>4
6.	(y6)	Math-2	-0.022	0.999	0.256	0.972	Math <et>4
7.	(y7)	Science	-.018	1.008	0.209	0.881	Math <et>4 Verbal <et>5
8.	(y8)	Read	-0.028	1.001	0.329	0.933	Verbal <et>5
9.	(y9)	Vocabulary	-0.037	0.998	0.437	0.914	Verbal <et>5
10.	(y10)	Write	-0.033	1.003	0.375	0.880	Math <et>4 Verbal <et>5
11.	(y11)	EDASP-1	-0.015	1.005	0.236	0.883	EDASP <et>6
12.	(y12)	EDASP-2	-0.039	0.997	0.467	0.906	EDASP <et>6
13.	(y13)	EDASP-3	-0.030	0.994	0.354	1.000	EDASP <et>6
14.	(y14)	Track	-0.049	0.984	0.584	1.014	Track <et>7
Senior Outcomes (based on 1982 data)							
15.	(y15)	Math-1	0.113	1.091	0.698	0.940	Math <et>8
16.	(y16)	Math-2	0.094	1.043	0.492	1.074	Math <et>8
17.	(y17)	Science	0.149	1.030	0.459	0.883	Math <et>8 Verbal <et>9
18.	(y18)	Read	0.200	1.068	0.687	0.972	Verbal <et>9
19.	(y19)	Vocabulary	0.328	1.069	0.930	0.869	Verbal <et>9
20.	(y20)	Write	0.261	1.009	0.744	0.797	Math <et>8 Verbal <et>9
21.	(y21)	EDASP-1	-0.039	1.012	0.287	0.847	EDASP <et>10
22.	(y22)	EDASP-2	0.010	0.942	0.519	0.876	EDASP <et>10
23.	(y23)	EDASP-3	.142	1.018	0.630	0.996	EDASP <et>10
24.	(y24)	TRACK	0.020	1.007	0.765	0.947	Track <et>11
25.	(y25)	Credits	-0.068	0.969	0.836	0.998	Credits <et>12
Post-Secondary Outcome Variables (based on 1984 data)							
26.	(y26)	POSTSEC-1	-0.040	0.996	0.514	0.910	POSTSEC <et>13
27.	(y27)	POSTSEC-2	-0.038	0.994	0.491	0.949	POSTSEC <et>13

Note. See Appendix 1 for a description of the variables. Y1-y14 and y25-y27 were standardized so as to have Mean=0, SD=1 across the total sample. Each senior variable (y15-y24) that was paired with a sophomore variable (y5-y14) was standardized using the total sample mean of the corresponding sophomore variable. The latent constructs $\langle et \rangle_1$-$\langle et \rangle_{13}$ associated with each measured variable are the ovals in Figure 1.

Table 2

Goodness of Fit Indexes for Alternative Models

Model	χ^2	df	RMSR	TLI	BBI
M1	1450	632	.024	.987	.981
M2	1343	625	.023	.989	.983
M3	1194	492	.022	.986	.984
M4	1490	631	.025	.986	.981
M5	1521	642	.028	.986	.980
M6	1320	595	.023	.9--	.9--
M7	1307	571	.023	.9--	.9--
M8[a]	1343	625	.023	.989	.983
M9[a]	1343	625	.023	.989	.983
M10	1343	625	.023	.989	.983
M0	76621	784	.406	.000	.000

Model Descriptions (also see Figure 1)

M1 All parameter estimates invariant except for latent mean differences.

M2 Model M1 with Variances and Covariances of Background constructs (SES, Black, Hisp) not invariant across groups (Fig. 1).

M3 No parameters invariant except measured variable intercepts.

M4 Model M2 with adjusted mean differences on senior and post-secondary latent constructs fixed to be zero.

M5 Model M2 with all factor loadings and uniquenesses for matching sophomore and senior measured variables constrained to be equal over time.

M6 Model M2 with the 30 path coefficients leading from SES, Black, and Hisp to each of the remaining 10 constructs being freed.

M7 Model M6 after freeing the 24 path coefficients leading from the 4 sophomore outcome constructs to the 6 senior and POSTSEC constructs.

M8[a] Model M2 with Post-secondary latent mean adjusted for only background and sophomore outcome constructs (Fig. 1).

M9[a] Model M8 with POSTSEC, senior Math, Verbal, and EDASP latent means adjusted for background variables, sophomore outcome variables, and senior Track and Credits (Fig. 1).

M10 Model M8 with adjusted POSTSEC, senior Math, Verbal, and EDASP latent mean differences fixed to be zero.

M0 Null model in which the fitted matrices were constrained to be diagonal for both groups (used in computing the TLI and BBI indices).

Note. RMSR=Root Mean Square Residual. TLI=Tucker-Lewis Index. BBI=Benter-Bonett Index. (See Marsh, Balla & McDonald, 1988, for further discussion of these indices).

[a] Models M8 and M9 are necessarily equivalent to M2 in their ability to fit the data. They differ in the posited ordering of the latent constructs so that the latent means are adjusted for different constructs in determining public/Catholic differences.

Table 3

Parameter Estimates For Model (M2) For Measured Variables

Measured Variable	Factor Loadings[a]		Uniqueness	Correlated Uniqueness[b]	Intercepts
1. SES-1	1		0.211		.516
2. SES-2	1.047		0.135		.542
3. Race--Black	1		0		-.226
4. Race--Hispanic	1		0		-.099
Sophomore Outcomes (based on 1980 data)					
5. Math-1	1		0.130		.354
6. Math-2	0.777		0.474	0.091	.275
7. Science[a]	0.246	0.702	0.374	0.146	.339
8. Read	1		0.294	0.067	.358
9. Vocabulary	1.003		0.284	0.103	.359
10. Write[a]	0.199	0.726	0.409	0.187	.330
11. Educ Aspir-1	1		0.666	0.154	.290
12. Educ Aspir-2	1.578		0.174	0.010	.451
13. Educ Aspir-3	1.326		0.416	0.085	.378
14. AcdTrack	1		0		.584
Senior Outcomes (based on 1982 data)					
15. Math-1	1		0.125		.682
16. Math-2	0.786		0.444		.537
17. Science[a]	0.201	0.683	0.379		.604
18. Read	1		0.286		.712
19. Vocabulary	1.002		0.274		.850
20. Write[a]	0.159	0.672	0.418		.701
21. Educ Aspir-1	1		0.618		.325
22. Educ Aspir-2	1.353		0.187		.507
23. Educ Aspir-3	1.369		0.323		.643
24. AcdTrack	1		0		.765
25. AcadCredit	1		0		.836
Post-Secondary Outcome Variables (based on 1984 data)					
26. Post-Sec-1	1		0.196		.510
27. Post-Sec-2	0.973		0.239		.496

Note. The numbers associated with each measured variable correspond to those for the measured variables in Table 1 (also see Appendix). Parameter estimates of 1 and 0 were fixed (i.e., not estimated). All estimated parameter values are statistically significant ($p < .05$).

[a] Just these measured variables are associated two latent construts. The loading associated with the math construct is presented first, followed by the loading associated with the verbal construct. [b] These eight a priori correlated uniquenesses were posited between the same measured variables administered on two different occasions.

Table 4

Parameter Estimates Relating Latent Constructs In Model (M2): Construct Variances and Covariances (above main diagonal) and Path Coefficients (below main diagonal

Latent Construct	1	2	3	4	5	6	7	8	9	10	11	12	13
1. SES	(.777)[a]	-.215[b]	-.177[b]	0	0	0	0	0	0	0	0	0	0
2. Black	0	(1.042)[a]	-.145[b]	0	0	0	0	0	0	0	0	0	0
3. Hisp	0	0	(1.018)[a]	0	0	0	0	0	0	0	0	0	0
4. Math	.354	-.204	-.179	(.650)	.436	.168	.260	0	0	0	0	0	0
5. Verbal	.339	-.211	-.192	0	(.487)	.141	.232	0	0	0	0	0	0
6. Educ Asp	.339	.086	.013[c]	0	0	(.242)	.172	0	0	0	0	0	0
7. Track	.296	.022[c]	-.046	0	0	0	(.901)	0	0	0	0	0	0
8. Math	.032	-.039	-.030	.887	.045	.182	.017[c]	(.145)	.078	.021	.049	.070	0
9. Verbal	.012[c]	-.043	-.026	.036[c]	.935	.081	.011[c]	0	(.096)	.010	.018	.028	0
10. Educ Asp	.081	.050	.014	.025[c]	.113	.665	.029	0	0	(.130)	.071	0	0
11. Track	.047	.052	.028	.071	.206	.424	.303	0	0	0	(.613)	.135	0
12. AcadCred	.013[c]	.021[c]	.026[c]	.216	.149	.467	.163	0	0	0	0	(.587)	0
13. Post-Sec	.079	-.009[c]	-.007[c]	-.022[c]	-.125[c]	-.053[c]	-.005[c]	.133[c]	.132[c]	.763	.052	.088	(.345)
Latent Construct Variances[a]	.777	1.042	1.018	.786	.621	.316	.960	.950	.746	.367	.988	.924	.757

Note. See Figure 1 and Table 3 for the relations between each latent construct and the corresponding measured variables. All parameters values of 0 were fixed (i.e., not estimated). Parameter estimates in parentheses are variance (for constructs 1-3) or residual variance (for constructs 4-13) estimates for each latent construct (from PSI). Parameter estimates above the main diagonal are covariance or residual covariances (from PSI). Parameter estimates below the main diagonal are path coefficients (from <BE>) leading from column variable to the row variable. Variances are variances of the latent constructs which will necessarily be larger than the residual variance estimates in the main diagonal for $\langle et \rangle_1$ - $\langle et \rangle_{13}$. All parameter estimates were constrained to be the same for both groups except for those noted in footnotes a and b to this table.

[a] Construct variance estimates are based on the Public data. The corresponding estimates for the public school group are .642, .412 .761. [b] Construct covariance estimates are based on the Public data. The corresponding estimates for the Catholic (in parentheses) and public school groups are -.215 (-.046), -.177 (-.080), -.145 (-.038). [c] These parameter estimates are not significantly different from zero at $p < .05$.

Table 5

Catholic/Public Differences on Latent Constructs (positive values indicate Public means > Catholic means) for selected models

Latent Constructs	Model M1 Adj Mean	SE	Unadj Mean	Model M2 Adj Mean	SE	Unadj Mean	Model M3 Adj Mean	SE	Unadj Mean
Background Constructs									
1. SES	-.560	.056	-.560	-.560	.052	-.560	-.571	.052	-.571
2. Black	.244	.060	.244	.244	.041	.244	.247	.041	.247
3. Hisp	.106	.060	.106	.106	.053	.106	.112	.053	.112
Sophomore Outcome Constructs									
4. Math	-.118	.054	-.385	-.118	.054	-.385	-.139	.051	-.418
5. Verbal	-.127	.047	-.389	-.127	.047	-.389	-.140	.047	-.411
6. Educ Asp	-.142	.034	-.310	-.142	.034	-.310	-.147	.032	-.320
7. Track	-.468	.058	-.633	-.468	.058	-.632	-.478	.060	-.646
Senior Outcome Constructs									
8. Math	-.112	.035	-.568	-.112	.035	-.568	-.119	.034	-.609
9. Verbal	-.085	.029	-.515	-.085	.029	-.515	-.086	.028	-.539
10. Educ Asp	-.057	.028	-.367	-.057	.028	-.367	-.058	.027	-.081
11. Track	-.304	.049	-.745	-.304	.049	-.745	-.304	.050	-.763
12. AcadCred	-.516	.048	-.904	-.516	.048	-.904	-.523	.051	-.928
Post-Secondary Outcome Constructs									
13. Post-Sec	-.037[a]	.044	-.550	-.037	.044	-.550	-.036	.045	-.571

Note. As shown in Figure 1 (model M2) Sophomore Outcomes are adjusted for Background variables, Senior Outcomes are adjusted for background variables and sophomore outcomes, and the Post-secondary outcome is adjusted for background variables, sophomore and senior outcomes. The size of mean differences may be evaluated in relation to standard errors (SE) or the variances of the latent constructs (see Table 4).

[a] This latent POSTSEC mean difference is adjusted for all background constructs and for all sophomore and senior outcome constructs. In an alternative formulation (Model M8) in which POSTSEC was corrected for only background and sophomore constructs, the latent mean difference was -.169 with a standard error of .047. This suggests that there is a significant public/Catholic difference in POSTSEC beyond what can be explained in terms of background and sophomore outcomes, but that this effect is largely mediated through senior outcomes.

Table 6

Tests of Interaction Effects: Path Coefficients Leading From SES, Black, Hisp, and Sophomore Math and Verbal Scores (and standard errors) estimated independently for Catholic (Cath) and public (pub) high school students.

Latent Construct		SES		Black		Hisp		Soph Math		Soph Verb	
		Cath	Pub	Cath	Pub	Cath	Pub	Cath	Pub	Cath	Pub
Sophomore Outcomes											
4.	Math	.191[a]	.365	-.162	-.204	-.141	-.180	0	0	0	0
	SE	.067	.019	.079	.015	.058	.015				
5.	Verb	.262	.343	-.133	-.213	-.102	-.197	0	0	0	0
	SE	.059	.017	.070	.013	.052	.013				
6.	EDASP	.294	.342	.086	.086	.034	.012	0	0	0	0
	SE	.043	.014	.050	.010	.037	.009				
7.	Track	.332	.293	.064	.020	.012	.050	0	0	0	0
	SE	.073	.020	.087	.012	.064	.016				
Senior Outcomes											
8.	Math	.053	.030	-.045	-.039	-.001	-.032	.941	.883	.001	.041
	SE	.052	.015	.055	.011	.040	.010	.083	.034	.097	.035
9.	Verbal	.009	.012	-.019	-.043	-.010	-.023	.035	.035	.865	.940
	SE	.042	.012	.044	.009	.032	.009	.067	.024	.079	.030
10.	Educ Asp	.092	.080	.059	.050	.022	.013	.066	.021	.082	.116
	SE	.040	.012	.042	.009	.031	.008	.065	.020	.077	.023
11.	Track	.130	.040	.027	.053	.049	.028	.095	.068	.106	.215
	SE	.071	.020	.074	.015	.054	.014	.115	.035	.135	.041
12.	Credit	-.028	.016	.021	.021	.029	.020	.206	.216	.195	.147
	SE	.070	.020	.073	.015	.054	.014	.113	.035	.133	.040
Post-secondary Outcomes											
13.	Post-Sec	-.028	.082	-.080	-.005	-.070	-.002	-.023	-.023	-.144	-.123
	SE	.068	.018	.066	.014	.048	.013	.134	.088	.154	.112

Note. In this model M7, path coefficients leading from background constructs and from sophomore outcome constructs were not constrained to be equal in the two groups (see Table 3). Selected path coefficients for each sample and their corresponding standard errors are presented.

[a] Differences between the public and Catholic path coefficients are statistically significant ($p < .05$). For purposes of these analyses, SEdiff was defined as $[(SE^2_{Cath}) + (SE^2_{pub})]^{1/2}$ (see Cohen and Cohen, 1985, p. 111, equation 3.6.11).

Appendix 1

Definition of Variables Considered

Social-Economic Status

1. SES-1980. [BYSES] 1980 Composite socioeconomic status based on father's occupation, father's education, mother's education, family income, and material possessions in the home (higher values reflect higher SES).

2. SES-1982. [FYSES] 1982 Composite socioeconomic status based on father's occupation, father's education, mother's education, family income, and material possessions in the home (higher values reflect higher SES).

Race/Ethnicity

3. Race--Black. [Race2] Ethnicity is Black. (1=yes, 0=no)

4. Race--Hispanic. [Race2] Ethnicity is Mexican, Cuban, Puerto Rican, or other Hispanic. (1=yes, 0=no)

Achievement Outcome Scores (sophomore and senior years)

5 & 15. Math achievement. [YBMTH1FS, FYMTH1FS] formula score for part 1 of the math tests for sophomore and senior years.

6 & 16. Math achievement. [YBMTH2FS, FYMTH2FS] formula score for part 2 of the math tests for sophomore and senior years.

7 & 17. Reading achievement. [YBREADFS, FYREADFS] reading test formula score for sophomore and senior years.

8 & 18. Vocabulary achievement. [YBVOCBFS, FYVOCBFS] vocabulary test formula score for sophomore and senior years.

9 & 19. Science achievement. [YBSCINFS, FYSCINFS] science test formula score for sophomore and senior years.

10 & 20. Writing achievement. [YBWRITFS, FYWRITFS] writing test formula score for sophomore and senior years.

Educational Aspirations (sophomore and senior years).

11. & 21. EDASP-1 [BB061G, FY76G] Item asking whether disappointed if do not graduate from college (higher scores reflect higher educational aspirations).

12. & 22. EDASP-2 [BB065, FY80] Item asking expected level of schooling (higher scores reflect higher educational aspirations).

13. & 23. EDASP-3 [BB067, FY82] Item asking lowest level of education would be satisfied with (higher scores reflect higher educational aspirations).

Academic Track (sophomore and senior years)

14. & 24. Track. [BB002; FY2] In 1980 and 1982 participated in academic track (1=yes, 0=no)

Academic Courses (senior year)

25. AcadCrd. [NEWBASE] In 1982 number of credits in six academic areas.

Post-Secondary Outcome Variables

26. Post-Sec. [PSESOC82, PSESFE83] Sum of variables indicating student was not a student (0), was a part-time student (1), or was a full-time student (2) at some form of post-secondary institution at each of two points in time during the first year after normal high school graduation.

27. Post-Sec. [PSESOC83, PSESFE84] Sum of variables indicating student was not a student (0), was a part-time student (1), or was a full-time student (2) at some form of post-secondary institution at each of two points in time during the second year after normal high school graduation.

Note. The number associated with each variable corresponds to the numbers in the boxes that represent measured variables in Figure 1. Values in brackets refer to variables names used on the HSB data file. Most outcome variables for the sophomore and senior years were paired and defined with parallel variables.

Appendix 2. The LISREL Multigroup Approach to Testing Group Differences in the Means of Latent Constructs.

The usual LISREL single group covariance analysis fits a sample covariance matrix for variables $\underline{x}$ and $\underline{y}$ with a theoretical matrix based on:

Insert Equations 1-3

The four parameter matrices (<LX>, <LY>, <TE>, and <TD>) are free to be estimated as factor loadings and uniquenesses, whereas the remaining four parameter matrices (<BE>, <GA>, <PH>, <PS>) are free to be estimated as latent partial regressions, latent variances and covariances, or latent residual variances and covariances. This model can be succinctly represented by the following path diagram (for which we gratefully acknowledge Jack Mc Ardle's influence).

Insert Figure 2

This usual LISREL approach does not incorporate regression intercept terms. Sorbom (1982) describes a technique for incorporating latent intercepts into a multigroup LISREL analysis, thus allowing for tests of latent intercepts which under appropriate circumstances can be interpretted as mean differences in the latent constructs. This approach can be illustrated with the following path diagram and definitions.

Insert Figure 3

Insert Equations 4-9

In this model, <al> and <mu> are latent and manifest regression intercepts. The truly estimatable parameter matrices with substantive interpretations are <mu>, <LY>, <TE>, <al>, <BE>, <PS>. These are embedded in the usual eight LISREL parameter matrices. In the present application (with 27 measured y-variables, 13 latent constructs, two groups and one fixed-X x-variables) these parameter matrices are:

a (27 x 13+1) matrix	[Equation 10]
a (27 x 27) matrix	[Equation 11]
a (1 x 1) matrix	[Equation 12]
a (1 x 1) matrix	[Equation 13]
a (13+1 x 13+1) matrix	[Equation 14]
a (13+1 x 1) matrix	[Equation 15]
a (1 x 1) matrix	[Equation 16]
a (13+1 x 13+1) matrix	[Equation 17]

In this formulation, the only role of $x \equiv 1$ is to allow the latent intercept terms (via regression of <et> on <xi> $\equiv x \equiv 1$) and to allow manifest intercept terms (via regression of y on <et> $\equiv$ <xi> $\equiv x \equiv 1$).

Such a structure is modelled in each of the groups. Supressing the 8s

and using the superscript g =1, 2, ... to indicate the groups we have

[Equation 18]

and under appropriate conditions of factorial invariance

[Equation 19]

An indeterminancy remains, as the data expectations are unaltered if replace

[Equation 20]

and

[Equation 21]

This can be removed by fixing each component of <al> in one group or the other to be zero. Each column of <LY> should have a (scaling) 1 and <BE> reflects a latent regression that is identified. Under these conditions, <al> reflects the latent variable mean differences. For example, suppose that in our application we have:

[Equation 22]

and

[Equation 23]

Then

[Equation 24]

So, subjects matched on <et> in the two groups will still differ on average in their $\langle ET \rangle_3$ scores by $\langle al \rangle_3^{(1)} - \langle al \rangle_3^{(2)}$. That is, this estimated difference reflects the mean $\langle ET \rangle_3$ group differences, adjusting for $\langle ET \rangle_2$. If <BE> is not invariant across groups, however, then subjects with the same $\langle ET \rangle_2$ in both groups will have

[Equation 25]

so that the difference in $\langle ET \rangle_3$ is a function of the level of $\langle ET \rangle_2$. This represents a group by $\langle ET \rangle_2$ interaction.

EQUATION 1:

$$\underset{\sim}{y} = \Lambda_y \underset{\sim}{\eta} + \underset{\sim}{\varepsilon}\ ;\ E\underset{\sim}{y} = E\underset{\sim}{\varepsilon} = \underset{\sim}{0};\ E\underset{\sim}{\eta} = \underset{\sim}{0};\ E\underset{\sim}{\varepsilon}\underset{\sim}{\varepsilon}' = \Theta_\varepsilon.$$

EQUATION 2:

$$\underset{\sim}{x} = \Lambda_x \underset{\sim}{\xi} + \underset{\sim}{\delta}\ ;\ E\underset{\sim}{x} = E\underset{\sim}{\delta} = \underset{\sim}{0};\ E\underset{\sim}{\xi} = \underset{\sim}{0}\ ;\ E\underset{\sim}{\delta}\underset{\sim}{\delta}' = \Theta_\delta.$$

EQUATION 3:

$$\underset{\sim}{\eta} = B\underset{\sim}{\eta} + \Gamma\underset{\sim}{\xi} + \underset{\sim}{\zeta}\ ;\ E\underset{\sim}{\zeta} = \underset{\sim}{0};\ E\underset{\sim}{\xi} = \underset{\sim}{0};\ E\underset{\sim}{\xi}\underset{\sim}{\xi}' = \Phi;\ E\underset{\sim}{\zeta}\underset{\sim}{\zeta}' = \Psi.$$

EQUATIONS 4 TO 9 (E4 TO E9):

E4: $x = \Lambda_x \xi + \delta = \xi$; where constants $\Lambda_x = 1$, $\Theta_\delta = \Phi = 0$ scalar one applied and "data" $x \equiv 1$ is supplied.

E5: $\underset{\sim}{\eta}^A = B^A \underset{\sim}{\eta}^A + \Gamma^A \xi + \underset{\sim}{\zeta}^A$; where

$$B^A = \begin{bmatrix} B & \underset{\sim}{0} \\ \underset{\sim}{0}' & 0 \end{bmatrix}_{14\times 14}$$

$$\underset{\sim}{\eta}^A = \begin{pmatrix} \underset{\sim}{\eta} \\ \eta^* \end{pmatrix}_{14\times 1}, \quad \eta^* \equiv \xi \equiv x \equiv 1$$

$$\Gamma^A = \begin{pmatrix} \underset{\sim}{\alpha} \\ 1 \end{pmatrix}_{14\times 1}$$

$$\underset{\sim}{\Psi}^A = \begin{bmatrix} \underset{\sim}{\Psi} & \underset{\sim}{0} \\ \underset{\sim}{0}' & 0 \end{bmatrix}_{14\times 14}$$

$$\underset{\sim}{\zeta}^A = \begin{pmatrix} \underset{\sim}{\zeta} \\ \zeta^* \end{pmatrix}_{14\times 1}, \quad \zeta^* \equiv 0$$

E6: $$\begin{pmatrix} \underset{\sim}{\eta} \\ \eta^* \end{pmatrix} = \begin{bmatrix} B & \underset{\sim}{0} \\ \underset{\sim}{0}' & 0 \end{bmatrix} \begin{pmatrix} \underset{\sim}{\eta} \\ \eta^* \end{pmatrix} + \begin{bmatrix} \underset{\sim}{\alpha} \\ 1 \end{bmatrix} \xi + \begin{pmatrix} \underset{\sim}{\zeta} \\ \zeta^* \end{pmatrix}$$

E7 : $\underset{\sim}{\eta} = B\underset{\sim}{\eta} + \underset{\sim}{\alpha}\xi + \underset{\sim}{\zeta} = \underset{\sim}{\alpha} + B\underset{\sim}{\eta} + \underset{\sim}{\zeta}$; $\xi \equiv x \equiv 1$.

E8 : $\eta^* = 1\xi + \zeta^* \equiv 1$; $E\zeta^*\zeta^{*\prime} = 0$

E9 : $\underset{\sim}{y} = \Lambda_y^A \underset{\sim}{\eta}^A + \underset{\sim}{\varepsilon} = [\Lambda_y | \underset{\sim}{\mu}]\underset{\sim}{\eta}^A + \underset{\sim}{\varepsilon}$

$= \underset{\sim}{\mu} + \Lambda_y \underset{\sim}{\eta} + \underset{\sim}{\varepsilon}$; $\eta^* \equiv \xi \equiv x \equiv 1$.

EQUATIONS 10 – 17 (E10 – E17):

E10 : $\Lambda_y^A = [\Lambda_y \; \underset{\sim}{\mu}]$ $27 \times (13+1)$

E11 : $\Theta_\varepsilon = \Theta_\varepsilon$ 27×27

E12 : $\Lambda_x = 1$ 1×1

E13 : $\Theta_\delta = 0$ 1×1

E14 : $B^A = \begin{bmatrix} B & \underset{\sim}{0} \\ \underset{\sim}{0}' & 0 \end{bmatrix}$ $(13+1) \times (13+1)$

E15 : $\Gamma = \begin{bmatrix} \underset{\sim}{\alpha} \\ 1 \end{bmatrix}$ $(13+1) \times 1$

E16 : $\underline{\Phi} = 0$ 1×1

E17 : $\Psi^A = \begin{bmatrix} \underline{\Psi} & \underset{\sim}{0} \\ \underset{\sim}{0}' & 0 \end{bmatrix}$ $(13+1) \times (13+1)$

EQUATION 18:

$$\underset{\sim}{y}^{(g)} = \underset{\sim}{\mu}^{(g)} + \Lambda_y^{(g)} \underset{\sim}{\eta}^{(g)} + \underset{\sim}{\varepsilon}^{(g)},$$

$$\underset{\sim}{\eta}^{(g)} = (I - B^{(g)})^{-1} \underset{\sim}{\alpha}^{(g)} + (I - B^{(g)})^{-1} \underset{\sim}{\zeta}^{(g)}$$

EQUATION 19:

$$E\,\underset{\sim}{y}^{(g)} = \underset{\sim}{\mu} + \Lambda_y \underset{\sim}{\eta}^{(g)}$$

$$= \underset{\sim}{\mu} + (I - B)^{-1} \underset{\sim}{\alpha}^{(g)}$$

EQUATION 20:

$$\underset{\sim}{\alpha}^{(g)} \quad \text{by} \quad \underset{\sim}{\alpha}^{(g)} - \underset{\sim}{\gamma}$$

EQUATION 21:

$$\underset{\sim}{\mu} \quad \text{by} \quad \underset{\sim}{\mu} + \Lambda_y (I - B)^{-1} \underset{\sim}{\gamma}$$

EQUATION 22:

$$\eta_3^{(1)} = \alpha_3^{(1)} + \beta_{32} \eta_2^{(1)} + \zeta_3^{(1)}$$

EQUATION 23:

$$\eta_3^{(2)} = \alpha_3^{(2)} + \beta_{32} \eta_2^{(2)} + \zeta_3^{(2)}$$

EQUATION 24:

$$\eta_3^{(1)} - \eta_3^{(2)} = \alpha_3^{(1)} - \alpha_3^{(2)} + \beta_{32}(\eta_2^{(1)} - \eta_2^{(2)}) + (\zeta_3^{(1)} - \zeta_3^{(2)}).$$

EQUATION 25:

$$E\eta_3^{(1)} - E\eta_3^{(2)} = \alpha_3^{(1)} - \alpha_3^{(2)} + \eta_2\left(\beta_{32}^{(1)} - \beta_{32}^{(2)}\right)$$

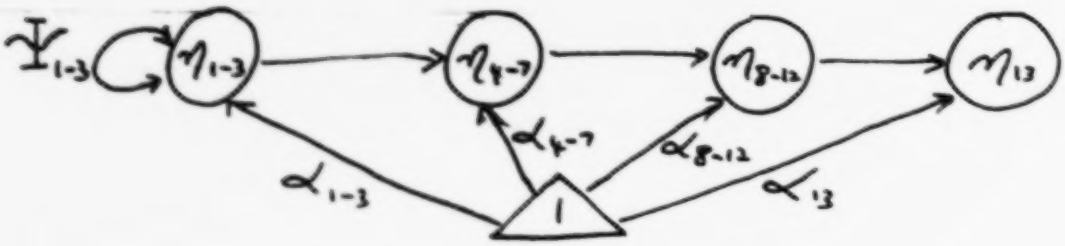
FIGURE CAPTIONS

Figure 1. Reduced path diagrams illustrating selected parameters from the structural portions of models M2, M8, and M9. All constructs are assumed to causally influence all constructs to their right as indicated by single-headed arrows, although only lag-1 arrows are actually presented. As presented in Table 1 $\langle et \rangle_1$-$\langle et \rangle_3$ are background constructs, $\langle et \rangle_4$-$\langle et \rangle_7$ are sophomore constructs, $\langle et \rangle_8$-$\langle et \rangle_{12}$ are senior constructs, and $\langle et \rangle_{13}$ is the post-secondary construct. The triangle represents the "constant" variable used by LISREL to estimate latent mean differences.

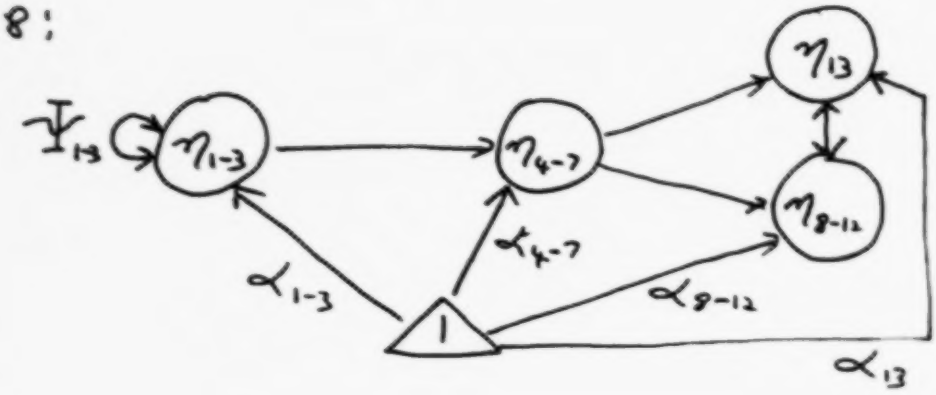
Figure 2 (Appendix 2). A condensed path-analytic representation of the general LISREL model. Single-headed arrows represent regressions and double-headed arrows represent covariances (or variances when self-referent).

Figure 3 (Appendix 2). A path-analytic representation of the general LISREL model that has been expanded to include regression intercept terms.

M2 :



M8:



M9:

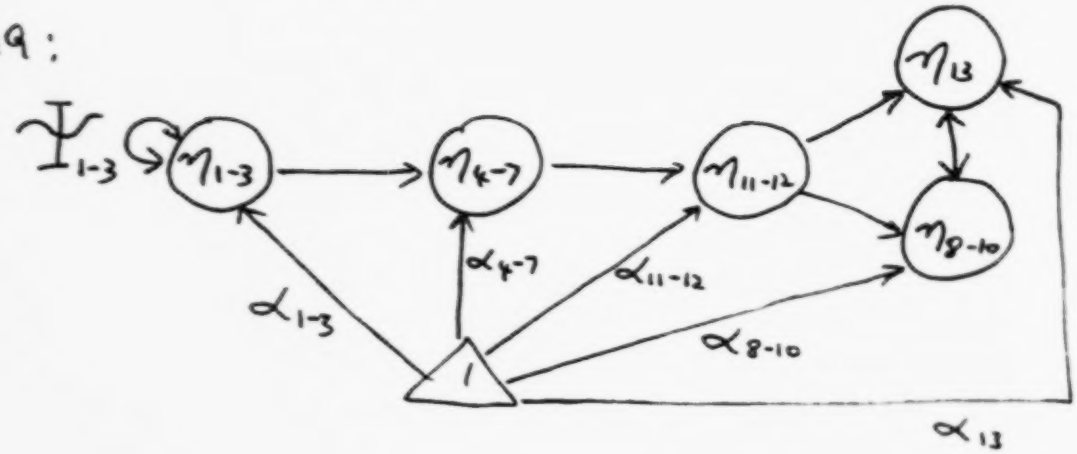


FIGURE 1. Reduced path diagrams for models M2, M8 and M9. See text for conventions.

FIGURE 2. LISREL PATH DIAGRAM.

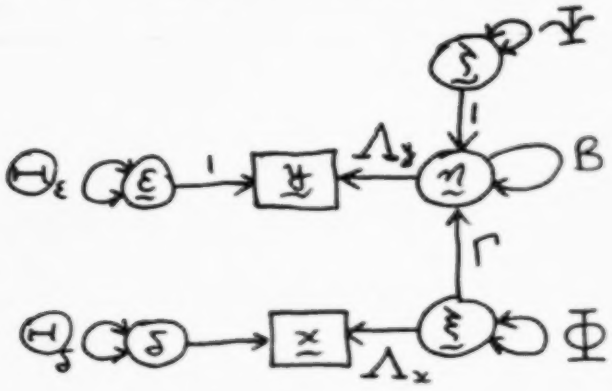


FIGURE 3. SINGLE GROUP WITH INTERCEPTS